2-D Transform-Domain Extended-Image Acquisition and Tracking Technique for Optical Pointing*

Haiping Tsou, Caroline Racho, and Tsun-Yee Yan

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

Abstract – An extended-image acquisition and tracking technique developed to enable precision pointing to a moving target through an imaging device is described in this paper. It is intended to be a fine-pointing scheme complimentary to the imager's "coarse" pointing control subsystem that is able to keep the target within the imager's field of view. This scheme compares the received extended-image covering more than one element of the detector array with a priorly established reference in the transform domain to estimate the target's movement. The received image is assumed to have each of its pixels corrupted by an independent additive white Gaussian noise. The coordinate of the target area is acquired and tracked, respectively, by an open-loop acquisition algorithm and a closed-loop tracking algorithm derived from the maximum likelihood criterion. This technique has many potential applications, e.g., free-space optical communications, where accurate and stabilized optical pointing is essential.

INTRODUCTION

Precision optical pointing is an essential capability of any laser-based system, especially when a laser signal is delivered to or received from a distant moving target. The pointing error caused by factors such as uncompensated platform motion and/or jitter, sensor noise and/or bias, and atmospheric propagation effects, such as image dancing, blurring, and scintillation [1], can significantly impair the system performance. In deep-space laser communications, an optical link may extend as long as several astronomy units (A.U., with 1 A.U. $\approx 149.6 \times 10^6$ kilometers), requiring high-power lasers with very narrow beam divergence for high-rate data transmission. The required pointing accuracy in this case is typically on the order of microradians on both ends of the optical link to ensure communication quality. Such a stringent requirement posts a real challenge in design of the pointing control system.

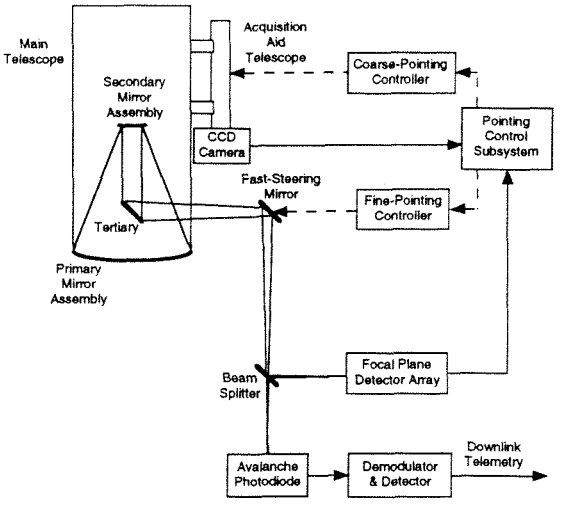


Figure 1: The optical terminal for laser communications.

A typical optical pointing control for deep-space applications is designed to be a two-level scheme. The coarse-pointing control is capable of keeping the target within the field of view of the optical detector by maintaining the telescope's pointing direction based on some pre-determined parameters, such as the predicted target trajectory, etc. On the other hand, the fine-pointing control is intended to track out any residual pointing error not being removed by the coarse-pointing. It is normally accomplished by driving a two-axis fast-steering mirror, as shown in Fig 1, such that the target image remains "fixed" on the focal plane detector array.

In this paper, a generic extended-image spatial acquisition and tracking technique developed to enable highly accurate and stable pointing to a moving target through an imaging device is presented. The detected image is assumed to be a randomly disturbed profile of a known target im-

* The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology under contract with the National Aeronautic and Space Administration.

age which covers more than one element of the imager's detector array. The proposed scheme estimates and tracks the target's movement based on the maximum likelihood criterion derived from the transform-domain correlation between the received image and a priorly established reference profile. It is assumed that the reference profile of the moving target exists or, at least, is able to be synthesized based on the current target position by using target information pre-stored and/or accumulated through continuous tracking[1]. In the following, we will briefly describe the analytical derivation, including the mathematical model for image representation, the effect of movements involving both translation and rotation in the discrete Fourier transform domain, the maximum likelihood estimator for spatial acquisition, and the image tracking loop developed for spatial tracking. Numerical results of sample scenarios to acquire a Sun-lit Earth beacon and to track a laser spot image modeled as a Gaussian pulse are also included, which demonstrate the capability to achieve sub-pixel resolution in high disturbance environments.

MOVEMENT OF RECEIVED IMAGE

The received image detected by an $M \times N$ focal plane array at time t_l, denoted as $r_l(m,n)$, can be represented by a sum of the source image, $s_l(m,n)$, and the random disturbance, $n_l(m,n)$, as follows

$$r_l(m,n) = s_l(m,n) + n_l(m,n), \tag{1}$$

where $m = 0, 1, \ldots, M-1$ and $n = 0, 1, \ldots, N-1$. With an additive white Gaussian random disturbance model, $n_l(m,n)$ is assumed to be an independent zero-mean Gaussian random variable with variance σ_l^2 for all m and n.

When, between t_l and t_{l+1}, the received image translates by the amount of x_l and y_l pixels along the x-axis and y-axis and rotates by an angle δ_l on the focal plane detector array, the resulting image at t_{l+1} is related to the previous image at t_l by

$$s_{l+1}(m,n) = s_l(m',n') + \epsilon_T(m',n') \tag{2}$$

where

$$m' \approx (m - x_l)\cos\delta_l + (n - y_l)\sin\delta_l \tag{3}$$
$$n' \approx -(m - x_l)\sin\delta_l + (n - y_l)\cos\delta_l \tag{4}$$

are the closest integers found to constitute the coordinate of a point in $s_l(m,n)$ at which the intensity is approximately equal to that at the corresponding point in $s_{l+1}(m,n)$, and $\epsilon_T(m',n')$ is the error introduced because of the fixed geometry of the detector array wherein each pixel has to be discretely represented. However, practically speaking, this error should be negligible as compared to the random disturbance from external sources and, therefore, is not considered in the analytical model to be discussed later.

In the (discrete Fourier) transform domain, it can be easily shown that the spatial-domain relationship stated in (2) becomes

$$\mathcal{S}_{l+1}(m,n) \approx \mathcal{S}_l(\alpha,\beta)e^{i\theta_{m,n,l}} \tag{5}$$

where $\mathcal{S}_{l+1}(m,n)$ is the transform-domain source image defined by

$$\sum_{p=0}^{M-1}\sum_{q=0}^{N-1} s_{l+1}(p,q)e^{-i2\pi\left(\frac{m}{M}p+\frac{n}{N}q\right)} \tag{6}$$

and, for the same reason mentioned above, α and β are the closest integers satisfying

$$\alpha \approx m\cos\delta_l + n\left(\frac{M}{N}\right)\sin\delta_l \tag{7}$$

$$\beta \approx n\cos\delta_l - m\left(\frac{N}{M}\right)\sin\delta_l \tag{8}$$

and

$$\theta_{m,n,l} = -2\pi\left(\frac{m}{M}x_l + \frac{n}{N}y_l\right) \tag{9}$$

is the phase introduced to the pixel (m,n) of the transform-domain image due to the translation of coordinate from t_l to t_{l+1}. It is interesting to note that the effects of translational and rotational movement are nicely separated in the transform domain, with the information about image translation contained in the phase and that about image rotation in the coordinate of the transform-domain image.

MAXIMUM LIKELIHOOD ESTIMATION

Based on the Gaussian assumption stated in (1), the maximum likelihood estimator will declare the estimated translation vector $(\hat{x}_l, \hat{y}_l)$ and the rotation angle $\hat{\delta}_l$ if

$$p(\vec{r}_{l+1} \mid \hat{x}_l, \hat{y}_l, \hat{\delta}_l; \vec{s}_l) = \max_{\{x_l, y_l, \delta_l\}} p(\vec{r}_{l+1} \mid x_l, y_l, \delta_l; \vec{s}_l) \tag{10}$$

where, for notational convenience, $\vec{r}_{l+1}$ and $\vec{s}_l$ are the vector representation[2] of the corresponding received image matrix $r_{l+1}(m,n)$ and the reference

[1] This assumption is generally valid when the relative motion between the target and the imager is largely confined on a two-dimensional plane during the acquisition and tracking.

[2] It is known as the lexicographic form in which the rows (or columns) of a matrix are concatenated sequentially to form a vector.

image matrix $s_l(m,n)$, respectively, and

$$p(\vec{r}_{l+1} \mid x_l, y_l, \delta_l; \vec{s}_l) = \frac{1}{\left(\sqrt{2\pi}\sigma_l\right)^{MN}} e^{-\frac{1}{2\sigma_l^2}\|\vec{r}_{l+1} - \mathbb{L}_{x_l,y_l,\delta_l}\{\vec{s}_l\}\|^2} \quad (11)$$

is the conditional probability density function of $\vec{r}_{l+1}$ given that the translation vector is (x_l, y_l), the rotation angle is δ_l, and the reference immediately before this movement is $\vec{s}_l$. Here, $\mathbb{L}_{x,y,\delta}\{\cdot\}$ is defined as a translation/rotation operator which moves the operand by a translation vector (x,y) and rotates it by an angle δ, and the notation $\|\vec{x}\|$ represents an $\mathcal{L}_2$ norm of the vector $\vec{x}$. The maximum likelihood criterion stated in (10) is equivalent to minimize the exponent in (11) over all possible (x_l, y_l) and δ_l, rendering the likelihood function to be maximized as

$$\mathrm{Re}\{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \mathcal{R}_{l+1}(m,n)\mathcal{S}_l^*(\alpha,\beta)e^{-i\theta_{m,n,l}}\} \quad (12)$$

where $\mathrm{Re}\{\cdot\}$ is the real part of a complex quantity.

As suggested by (10), the translation vector and rotation angle should be jointly estimated, which inevitably requires a massive parallel search through every possible coordinate shift and rotation angle since the information regarding the rotational movement is embedded in the coordinate of the transform-domain image. However, in reality, this problem can be greatly alleviated by breaking the one-shot joint estimation into two steps, first estimating the translation vector and then searching for the rotation angle with the help from information of the estimated translation. The reason behind this two-step approach is that, in most cases we are interested, the rotational movement has mostly been compensated for and, therefore, any residual is much slower than the transitional movement. A reasonably good convergence is normally achievable after repeating this process several times. In the following, we will concentrate on the algorithm of estimating the translation vector.

By taking the partial derivatives of the likelihood function in (12) with respect to x_l and y_l and equating them to zero, we have

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} m|\Omega_l(m,n)|\sin(\xi_{m,n,l} - \theta_{m,n,l}) = 0 \quad (13)$$

$$\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} n|\Omega_l(m,n)|\sin(\xi_{m,n,l} - \theta_{m,n,l}) = 0 \quad (14)$$

which form a set of simultaneous nonlinear equations to be solved for the maximum likelihood estimates of x_l and y_l. Here,

$$\Omega_l(m,n) \triangleq \mathcal{R}_{l+1}(m,n)\mathcal{S}_l^*(\alpha,\beta) \quad (15)$$

has magnitude $|\Omega_l(m,n)|$ and phase $\xi_{m,n,l}$. This set of equations provides the optimal spatial acquisition scheme for a rotation-invariant movement. When the image is close to being acquired, the phase differences, $(\xi_{m,n,l} - \theta_{m,n,l})$, are small and the approximation of $\sin(x) \approx x$ can be applied to (13) and (14), rendering a suboptimal linear estimator of which the computation complexity is greatly reduced.

IMAGE TRACKING LOOP

A closed-loop image tracking algorithm motivated by the same maximum likelihood criterion can also be developed for continuous tracking of a translational movement. In this case, it is the correlation between the transform-domain received image $\mathcal{R}_{l+1}(m,n)$ and the estimated transform-domain reference image

$$\hat{\mathcal{S}}_{l+1}(m,n) = \mathcal{S}_l(m,n)e^{i\hat{\theta}_{m,n,l}} \quad (16)$$

established from $\hat{\theta}_{m,n,l}$, the estimation of (9) of previous iteration, to be continuously monitored. It is found that, very similar to (12), the likelihood function becomes the average of $\mathrm{Re}\{C_{l+1}(m,n)\}$ over the entire detector array, where $C_{l+1}(m,n)$ is the pixel-wise product of $\mathcal{R}_{l+1}(m,n)$ and the conjugate of (16). This pixel-wise product can be shown to be a sum of zero-mean random disturbance associated with $n_l(m,n)$ in (1) and a deterministic part given by $|\mathcal{S}_l(m,n)|^2 e^{i\phi_{m,n,l}}$, where the estimation error

$$\begin{aligned}\phi_{m,n,l} &= \theta_{m,n,l} - \hat{\theta}_{m,n,l} \\ &= -2\pi\left[\frac{m}{M}(x_l - \hat{x}_l) + \frac{n}{N}(y_l - \hat{y}_l)\right] \\ &\triangleq -2\pi\left(\frac{m}{M}\Delta_x + \frac{n}{N}\Delta_y\right)\end{aligned} \quad (17)$$

with Δ_x and Δ_y being the associated errors in the estimated coordinate.

In deriving the tracking algorithm to continuously update the estimates, two simultaneous loop feedback signals, denoted as ε_x and ε_y, are formed as the partial derivatives of the likelihood function with respect to Δ_x and Δ_y, rendering

$$\begin{aligned}\varepsilon_x = &\frac{2\pi}{M}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} m|\mathcal{S}_l(m,n)|^2\sin(\phi_{m,n,l}) \\ &+ \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\mathcal{N}_{l,eff}^{(x)}(m,n)\end{aligned} \quad (18)$$

$$\begin{aligned}\varepsilon_y = &\frac{2\pi}{N}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} n|\mathcal{S}_l(m,n)|^2\sin(\phi_{m,n,l}) \\ &+ \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\mathcal{N}_{l,eff}^{(y)}(m,n)\end{aligned} \quad (19)$$

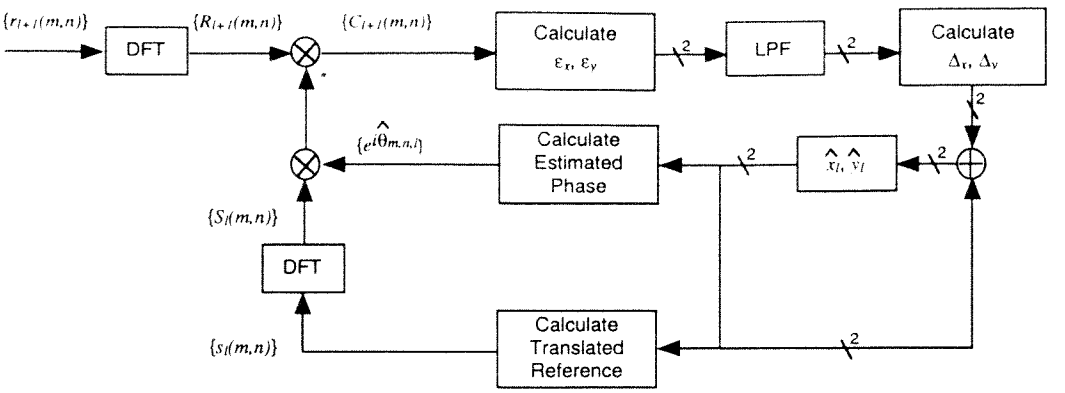


Figure 2: The Extended-Image Tracking Loop.

where $\mathcal{N}_{l,eff}^{(x)}(m,n)$ and $\mathcal{N}_{l,eff}^{(y)}(m,n)$ are the effective noises in the loop operation. Equations (18) and (19) characterize the relationship between the estimate errors and the loop feedback signals. However, to solve for Δ_x and Δ_y from these nonlinear equations can be a quite challenging task. With a reasonable assumption that the phase error $\phi_{m,n,l}$ remains small during the tracking mode, one can substitute (17) for $\sin(\phi_{m,n,l})$ in (18) and (19). The resulting linearized simultaneous equations can be easily solved, yielding

$$\Delta_x = \frac{C_n \mathbb{E}[\varepsilon_x] - C_{mn}\mathbb{E}[\varepsilon_y]}{C_m C_n - C_{mn}^2} \tag{20}$$

$$\Delta_y = \frac{C_{mn} \mathbb{E}[\varepsilon_x] - C_m \mathbb{E}[\varepsilon_y]}{C_{mn}^2 - C_m C_n} \tag{21}$$

where $\mathbb{E}[\cdot]$ denotes the statistical expectation and

$$C_m \triangleq \frac{4\pi^2}{M^2} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} m^2 |\mathcal{S}_l(m,n)|^2$$

$$C_n \triangleq \frac{4\pi^2}{N^2} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} n^2 |\mathcal{S}_l(m,n)|^2$$

$$C_{mn} \triangleq \frac{4\pi^2}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} mn |\mathcal{S}_l(m,n)|^2$$

are coefficients that can be calculated from the transform-domain reference image of the previous iteration at t_l.

The resulting image tracking loop is depicted in Fig. 2, in which the time averages performed by low-pass filtering replace the statistical averages found in (20) and (21). The calculated Δ_x and Δ_y are then used to update the movement estimates through an accumulator, such that

$$\hat{x}_{l+1} = \hat{x}_l + \Delta_x \tag{22}$$

$$\hat{y}_{l+1} = \hat{y}_l + \Delta_y \tag{23}$$

The updated accumulator contents will be used to calculate the estimate $\hat{\theta}_{m,n,l+1}$ and prepare the translated reference image for the next loop iteration at t_{l+1}.

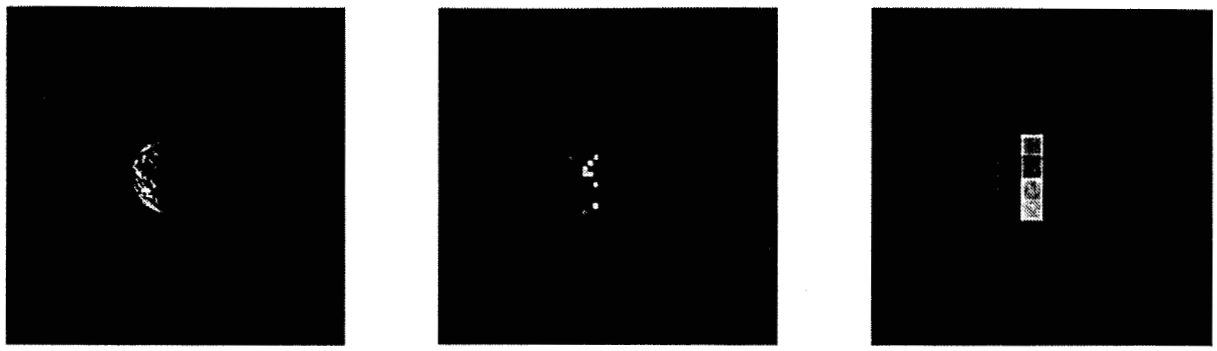
NUMERICAL RESULTS

The presented technique can be applied to many areas, especially for those with an asymmetric image profile and/or a very low signal-to-noise ratio (SNR) where the conventional spatial-domain centroid-type algorithms fail. Here, two numerical simulations for deep-space optical communications are presented to demonstrate its capability to achieve sub-pixel accuracy in estimating and tracking rotation-invariant movements in high disturbance environments.

First, the acquisition of an Earth beacon from an optical transceiver onboard spacecraft has been simulated by using Sun-lit Earth images shown in Fig. 3. The reference image is compressed from a 256 × 256 original taken by the Galileo spacecraft during its mission to Jupiter. The received image is assumed to be detected by a 16×16 array and corrupted by additive white Gaussian disturbances such that the average SNR defined as

$$\left(\frac{S}{N}\right) \triangleq \frac{1}{\sqrt{MN}} \frac{\| \vec{s}_l \|}{\sigma_l} \tag{24}$$

is unity in this simulation. The acquisition process suggested by (13) and (14) requires $M \times N$ terms of summation for each of the simultaneous equations. The actual size of computation can be reduced by first forming partial sums of these equations and then gradually including more terms in these partial sums until estimated results converge. Table 1 shows the convergence as the size of computation, specified by M_c and N_c, increases. The final results give a sub-pixel accuracy of less than 2% and 1% in the respective direction of the original reference image.



(a) (b) (c)

Figure 3: The Extended-Source Images: (a) original image of size 256 × 256, (b) reference image of size 64 × 64, and (c) received image by a detector array of size 16 × 16.

Indices		Est. Coordinate		Real Coordinate	
M_c	N_c	Row	Col.	Row	Col.
4	4	3.75	-3.14	3.625	-2.82
6	6	3.55	-2.84		
8	8	3.55	-2.81		

Table 1: Estimated Ground Station Coordinate

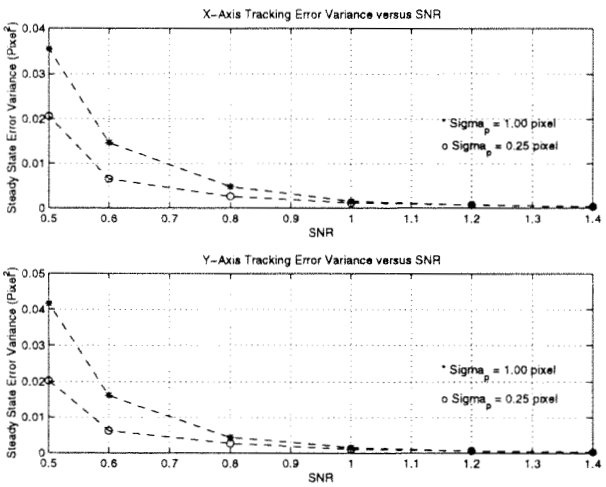


Figure 4: The Error Variance for Image Tracking.

The image tracking loop has been simulated for the downlink laser beam tracking. In this simulation, the intensity profile of the tracked laser spot is assumed to be a two-dimensional Gaussian pulse with its spread specified by the standard deviation, denoted as σ_p, of the Gaussian density function. It is further assumed that a 4×4 detector array is used to capture the laser spot image, rendering only about two-thirds of the laser power being collected when $\sigma_p = 1.0$ pixel versus over 99% of power when $\sigma_p = 0.25$ pixel as the center of the laser spot is initially offset by one pixel in both x- and y-direction on the detector array. Figure 4 shows the tracking error variance for different σ_p versus various SNRs defined in (24). It appears that tracking of a spread beam (e.g., $\sigma_p = 1.0$ pixel versus $\sigma_p = 0.25$ pixel) suffers no significant degradation except in the extremely low SNR region, which proves to be a key advantage over the spatial-domain centroid-type tracking algorithms. The numerical results indicate that an exponentially decay of error variance as SNR increases, with the standard deviation of the tracking error being 3.32×10^{-2} pixel for $\sigma_p = 0.25$ pixel and 3.87×10^{-2} pixel for $\sigma_p = 1$ pixel when SNR=1.

CONCLUSION

Optical pointing to a moving target has been an essential function in many laser-based applications. This paper describes a generic maximum-likelihood-based transform-domain correlation-type extended-image acquisition and tracking technique which has been demonstrated through numerical simulations to be able to achieve sub-pixel accuracy in estimating and tracking relative translation movement between the target and the imager in high disturbance environments. With an additional rotation angle estimation assisted by the information from the translation vector estimation to compensate for possible image rotation, this scheme can provide a cost-effective way to enhance pointing accuracy without putting too much extra burden on the imager's pointing control mechanism.

REFERENCES

[1] R. M. Gagliardi and S. Karp, *Optical Communications*, John Wiley & Sons, New York, 2 ed., 1995.